



ArchiveSIC, Archive Ouverte en Sciences de l'Information et de la Communication : rôle, fonctionnement et usage.

Gabriel Gallezot

► To cite this version:

Gabriel Gallezot. ArchiveSIC, Archive Ouverte en Sciences de l'Information et de la Communication : rôle, fonctionnement et usage.. 2003, pp.XX-XX. sic_00000600

HAL Id: sic_00000600

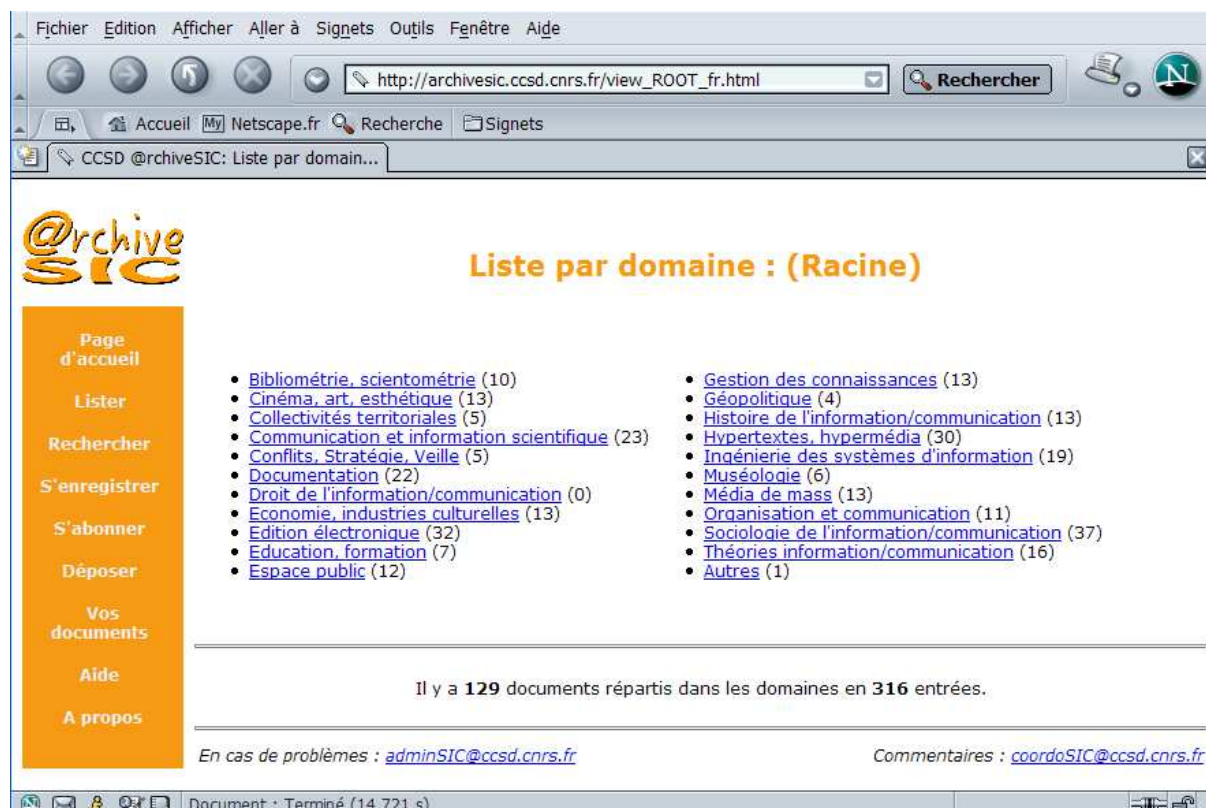
https://archivesic.ccsd.cnrs.fr/sic_00000600

Submitted on 18 Sep 2003

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ArchiveSIC, Archive Ouverte en Sciences de l'Information et de la Communication : rôle, fonctionnement et usage.



La création des archives ouvertes est née en réaction au paradoxe de l'édition scientifique : alors que le chercheur offre ces résultats, la communauté scientifique trouve un certain nombre d'obstacles pour lire ces publications. Pourtant la majorité des chercheurs accèdent aujourd'hui à Internet et pourraient se saisir de cette opportunité pour partager plus facilement les résultats de leurs travaux. De plus, l'inflation des écrits et la nécessité d'une validation dans un temps plus court pour certains résultats (souvent expérimentaux) se heurtent aux limites du modèle traditionnel de publication.

Le développement des archives ouvertes se fonde sur un processus d'auto-archivage des publications scientifiques par les auteurs. Tout en continuant à soumettre leurs articles dans des revues ou colloques, ceux-ci déposent parallèlement des versions pré-publiées (avant l'évaluation scientifique d'une revue) ou post-publiées (après l'évaluation scientifique d'une revue) dans des bases de données hypertextuelles lisibles de tous, les archives ouvertes ; l'objectif étant de maximiser la visibilité de ces articles et travaux à l'heure des réseaux.

Des revues SIC trop méconnues et dispersées

L'initiative d'@rchiveSIC est partie du constat alarmant dressé par une enquête récente du Ministère de la Recherche en France concernant les revues en Sciences de l'Information et de la Communication (SIC). Sur une liste de 263 revues recensées en SIC au plan mondial, 17 seulement sont connues d'au moins un chercheur sur deux. Quand on passe au plan national, 12 seulement des 40 revues françaises citées en SIC sont connues (mais pas forcément lues) d'au moins un chercheur sur deux.

En réponse à cette situation, un groupe de chercheurs¹ a proposé la création d'une Archive ouverte en SIC : @rchiveSIC. Il s'agissait de donner aux articles la lisibilité qu'ils méritent, faciliter la recherche et l'accès au

¹ Gallezot (G.), Chartron (G.), Noyer (J.-M.), "Une archive ouverte des publications en InfoCom", *colloque "Place et enjeux des revues pour la recherche en infoCom"*, SFSIC, Nice, 25-26 Mars 2002, <http://archivesic.ccsd.cnrs.fr/documents/archives0/00/00/00/25/index_fr.html>

savoir contenu dans les articles, relever le défi de la lutte contre la « babelisation » de la discipline tout en conservant la diversité qui en fait sa richesse.

Le choix Eprint.org

Pour réaliser cette archive, l'équipe fondatrice s'appuie sur les compétences techniques du CCSD (Centre pour la Communication Scientifique Directe) du CNRS qui maintient déjà des archives ouvertes : thèses-EN-ligne (<http://tel.ccsd.cnrs.fr/>) et le site miroir d'ArXiv (<http://ccsd.cnrs.fr/arXiv.html>).

L'archive est construite à partir du logiciel Eprint.org² (version 1.1), logiciel libre et gratuit, librement modifiable et déjà utilisé par plus d'une soixantaine de site (mars 2003). Sa compatibilité OAI rend de plus le projet interopérable avec d'autres archives. Partant des spécificités de communauté scientifique, Eprint.org est adapté et certaines fonctions ajoutées dans le cadre de la réalisation d'@rchiveSIC :

- Simplification du processus de dépôt en 3 écrans au lieu d'une dizaine,
- Possibilité de multi-dépôts, tout en conservant la « thématisation » de ceux-ci,
- Ouverture des dépôts à des formats fortement utilisés (rtf, .doc, .zip), vérification des en-têtes des formats, générateur automatique de format HTML et PDF,
- Modération partagée,
- Recherche full-text,
- Version bilingue (la version 1.1 d'Eprint.org n'offre pas en standard cette possibilité),
- Statistique de dépôt.

A noter que depuis ce choix technique, d'autres logiciels ou plates-formes techniques sont depuis disponibles et proposent des solutions « clés en main » d'auto-archivage (Dspace du MIT³, CDSware du CERN⁴ et HAL du CCSD⁵).

Un fonctionnement simplifié

L'archive est ouverte uniquement à la communauté des enseignants et chercheurs francophones en SIC. Après s'être enregistré, l'auteur est invité à déposer son texte. Trois types de documents sont acceptés (les statuts « working paper », « soumis à validation », « publié ou en cours de publication »), 20 thématiques proposées et 4 formats de dépôt principaux (RTF, DOC, PDF, HTML) acceptés (les formats PDF et HTML sont automatiquement générés si le dépôt s'est effectué dans d'autres formats, assurant ainsi un minimum de lisibilité, voir de pérennité).

Chaque thématique a son modérateur. Ce dernier vérifie uniquement que les articles déposés correspondent à la vocation de l'archive et assure un filtrage minimum (non une validation scientifique du contenu). Cette modération partagée élimine de surcroît toute velléité éditoriale d'un modérateur puisqu'un texte peut être classé dans plusieurs thématiques et donc être validé ou invalidé par un autre modérateur.

Bien que l'archive s'adresse principalement aux francophones, l'interface est volontairement bilingue pour permettre une recherche mondiale des textes contenus dans l'archive. De la même façon, la langue de dépôt n'est pas restreinte au seul français, mais ouverte à toutes les langues dans lesquelles un auteur de la communauté a pu communiquer (principalement : anglais, allemand, espagnol, italien). Parallèlement, pour tenir compte des différents niveaux d'appropriation des NTIC par les utilisateurs, les interfaces de dépôt et d'interrogation ont été simplifiées au maximum. A noter que chaque utilisateur de l'archive peut s'abonner aux thématiques scientifiques qui servent au classement des textes et choisir la périodicité de l'alerte des nouveaux textes.

Les métadonnées sont présentes uniquement dans la notice bibliographique du dépôt⁶. Elles respectent le standard OAI, qui s'exprime principalement par un balisage XML des éléments du *Dublin Core*. Ainsi, en respectant ce standard les textes présents dans l'archive peuvent être interconnectés à d'autres bases compatibles

² <http://www.eprints.org/>

³ Durable Digital Repository , <http://www.dspace.org/>

⁴ CERN Document Server Software, <http://cdsware.cern.ch/>

⁵ Hyper article en ligne, <http://hal.ccsd.cnrs.fr/>

⁶ Une feuille de style (DTD XML) aurait pu affiner le référencement du texte, mais aurait considérablement bloqué la dynamique de dépôt

OAI. D'autre part, l'ensemble des notices se trouve dans un répertoire accessible aux moteurs de recherche, les métadonnées OAI permettent globalement une bonne indexation⁷.

Des débuts encourageants

@rchiveSIC est disponible depuis juin 2002, elle compte 235 inscrits avec une progression d'inscription de 18 comptes par mois (déposants ou abonnement thématique). Elle contient à ce jour 125 documents (avec en moyenne 11 documents supplémentaires par mois) et comptabilise plus de 40 000 visites.

Le frein principal au dépôt semble relever du rapport « temps/utilité ». Des déposants accaparés par d'autres charges de travail ne trouvent, actuellement pas d'utilité immédiate à ce projet, le frein n'est donc pas d'ordre technologique mais culturel. Comment étendre l'usage de l'archive dans les pratiques des chercheurs en SIC ? Plusieurs stratégies sont actuellement envisagées en partant de l'hypothèse que l'appropriation doit permettre un gain d'efficacité pour chacun :

- *L'archive comme support de diffusion des actes de colloques et conférences*
- *L'archive comme diffusion parallèle des revues*
- *L'archive comme adjuvant aux articles édités*

Bibliographie

Bougnoux (Daniel), Rasse (Paul), "Les revues en SIC ", La lettre d'Inforcom, n°59, juin 2001.

Charnay (Daniel), « Le libre accès : quels enjeux pour la communauté scientifique ? », In Libre accès à l'information scientifique et technique : État de l'art et perspectives, Paris, 23-24 janvier 2003 - Carré des Sciences - Ministère de la Recherche, <http://www.inist.fr/openaccess/transparentes/charnay/index_fichiers/frame.htm>

Chartron G. (dir) Les mutations de l'édition de recherche, Internet et les chercheurs, Cercle de la Librairie, 2002.

Crawford (S. Y.), Hurd (J. M.), Weller (A.C.), From Print to Electronic, The Transformation of Scientific Communication, Medford, Information today Inc (Asis monograph series), Washington, USA, 1996.

De la Vega (Josette F.), La communication scientifique à l'épreuve de l'Internet, l'émergence d'un nouveau modèle, Presses de l'Enssib, 2000.

Dobratz (Susanne), Matthaei (Birgit), "Open Archives Activities and Experiences in Europe, An overview by the Open Archives Forum, D-Lib Magazine, Vol.9, n°1, janvier 2003, <<http://www.dlib.org/dlib/january03/dobratz/01dobratz.html>>

Gallezot (G.), Chartron (G.), Noyer (J.-M.), "Une archive ouverte des publications en InfoCom", colloque "Place et enjeux des revues pour la recherche en infoCom", SFSIC, Nice, 25-26 Mars 2002, <http://archives.ic.ccsd.cnrs.fr/documents/archives0/00/00/00/25/index_fr.html>

Ginsparg (Paul), "First Steps Towards Electronic Research Communication", Computers in Physics, vol.8, n°4, 1994, repris dans Solaris n°3, <<http://www.info.unicaen.fr/bnum/jelec/Solaris/d03/3ginspar.html>>.

Harnad (Stevan), "Post-Gutenberg Galaxy: The Fourth Revolution in the Means of Production of Knowledge", Public Access Computer Systems Review 2 (1), 1991, pp.39-53.
Voir également l'ensemble des publications de Stevan Harnad à l'adresse <<http://cogsci.soton.ac.uk/~harnad/intpub.html>>.

Harnad (Stevan), "For Whom the Gate Tolls? How and Why to Free the Refereed Research Literature Online Through Author/Institution Self-Archiving", 2001, <<http://cogprints.soton.ac.uk/documents/disk0/00/00/16/39/index.html>>

Jeannin (Philippe), Enquête sur les périodiques en Sciences de l'Information et de la Communication, 2001, <http://www.iut-tarbes.fr/enquete/info_comm/>

Noyer (Jean-Max). " L'économie du savoir ", Solaris n°1, 1994, <<http://www.info.unicaen.fr/bnum/jelec/Solaris/d01>>

⁷ Prise en compte du *Dublin core* et *Link popularity* par le standard OAI

Van de Sompel (H.), Lagoze (C). "The Santa Fe Convention of the Open Archives Initiative." D-Lib Magazine, 2000,
<<http://www.dlib.org/dlib/februaru00/vandesompel-oai/vandesompel-oai.html>>